Clustering of respiratory symptoms in the general population

Daniil Lisik[1], Lowie Vanfleteren[2,3], Rani Basna[1], Hannu Kankaanranta[1,4,5], Bright I Nwaru[1,6], Helena Backman[7]

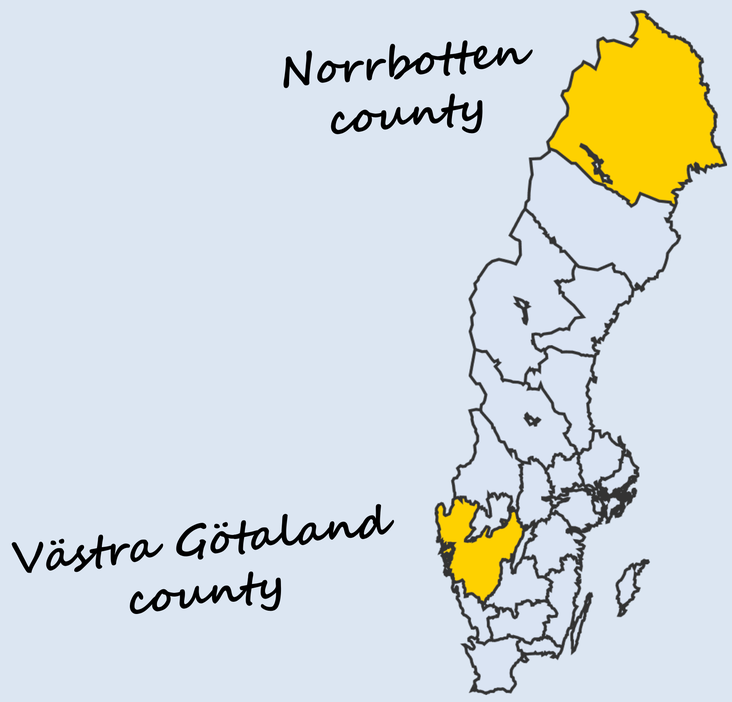
Contact: daniil.lisik@gu.se

Aim

To perform computational phenotyping of respiratory symptoms in a population-representative selection of adults.

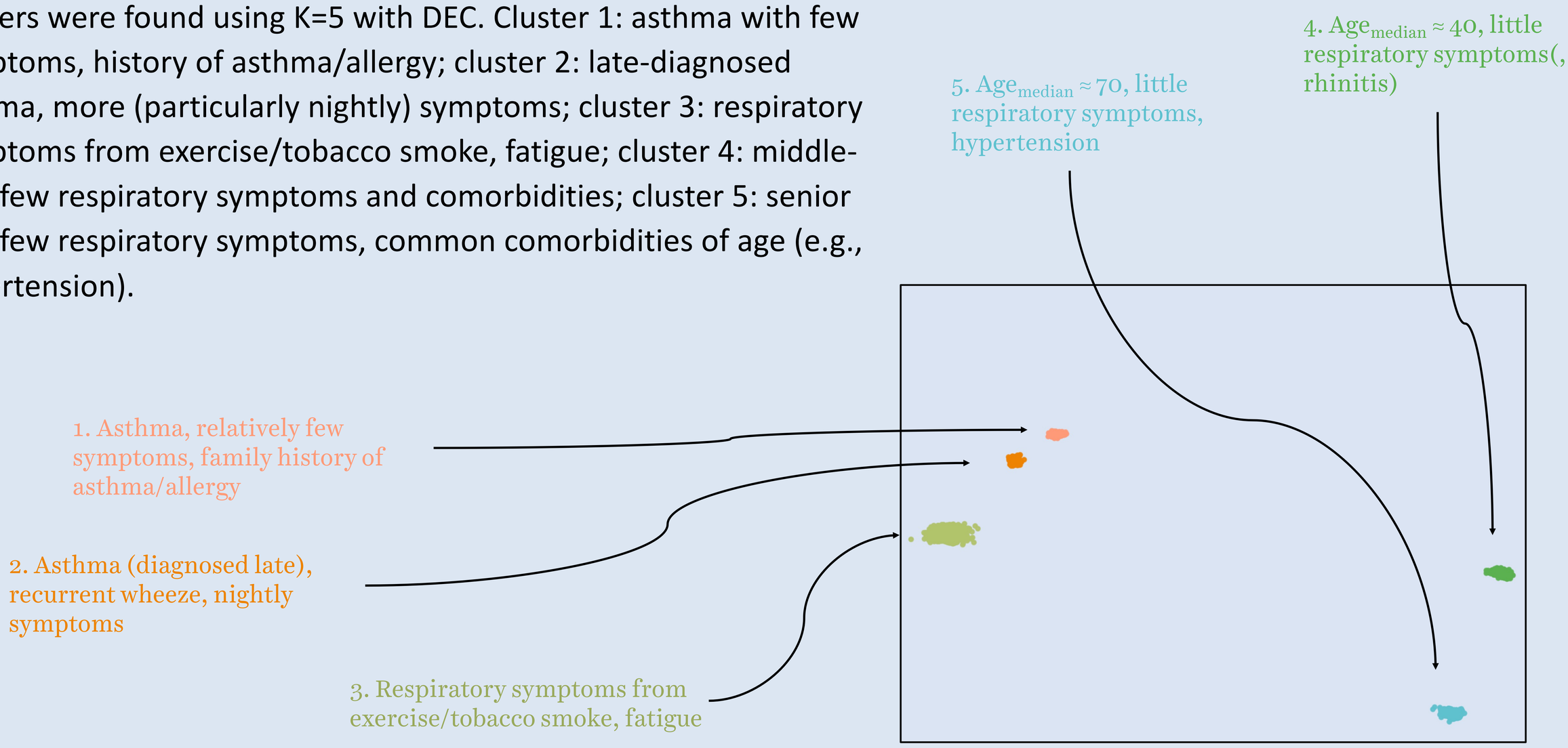
Methods

Survey data on respiratory health/symptoms were gathered from population-based cohorts in Northern and Western Sweden. The initial selection of 50 variables was reduced to 10 using principal component analysis (PCA). Four algorithms (K-means, deep embedded clustering (DEC), Dip-based deep embedded clustering (DipDeck), and Kingdra) were used for clustering. Internal validation was assessed with Silhouette score.



Results

The study population consisted of 61,530 adults. The strongest clusters were found using K=5 with DEC. Cluster 1: asthma with few symptoms, history of asthma/allergy; cluster 2: late-diagnosed asthma, more (particularly nightly) symptoms; cluster 3: respiratory symptoms from exercise/tobacco smoke, fatigue; cluster 4: middle-age, few respiratory symptoms and comorbidities; cluster 5: senior age, few respiratory symptoms, common comorbidities of age (e.g., hypertension).



Conclusion

Clustering of self-reported respiratory symptoms has potential, but involves challenges for identifying novel phenotypes. Further optimization of embedding and clustering models is needed to improve robustness and resolution in clustering.

[1] Krefting Research Centre, University of Gothenburg, Gothenburg, Sweden
[2] COPD Center, Department of Respiratory Medicine and Allergology, Sahlgrenska University Hospital, Gothenburg, Sweden
[3] Department of Internal Medicine and Clinical Nutrition, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden
[4] Tampere University Respiratory Research Group, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland
[5] Department of Respiratory Medicine, Seinäjoki Central Hospital, Seinäjoki, Finland
[6] Wallenberg Centre for Molecular and Translational Medicine, University of Gothenburg, Sweden
[7] Department of Public Health and Clinical Medicine, Section of Sustainable Health/the OLIN Unit, Umeå University, Umeå, Sweden